On Pseudocodewords and Improved Union Bound of Linear Programming Decoding of HDPC Codes

Ohad Gidon and Yair Be'ery *Senior Member*, *IEEE*

Tel Aviv University, School of Electrical Engineering

Ramat Aviv 69978, ISRAEL

Email: ohadgodo@post.tau.ac.il, ybeery@eng.tau.ac.il

Abstract

In this paper, we present an improved union bound on the Linear Programming (LP) decoding performance of the binary linear codes transmitted over an additive white Gaussian noise channels. The bounding technique is based on the second-order of Bonferroni-type inequality in probability theory, and it is minimized by Prim's minimum spanning tree algorithm. The bound calculation needs the fundamental cone generators of a given parity-check matrix rather than only their weight spectrum, but involves relatively low computational complexity. It is targeted to high-density parity-check codes, where the number of their generators is extremely large and these generators are spread densely in the Euclidean space. We explore the generator density and make a comparison between different parity-check matrix representations. That density effects on the improvement of the proposed bound over the conventional LP union bound. The paper also presents a complete pseudo-weight distribution of the fundamental cone generators for the BCH[31,21,5] code.

arXiv:1203.1850v1 [cs.IT] 8 Mar 2012

I. INTRODUCTION

THE calculation of error probability for Linear Programming (LP) decoding of Binary Phase-Shift Keying (BPSK) modulated binary codes is often a complex task. This is mainly due to the complexity of LP Voronoi or decision regions [1] [2]. The probability of correct decision in an Additive White Gaussian Noise (AWGN) channel, can be obtained by integrating a multidimensional Gaussian distribution over the decision region of the transmitted codeword (CW).

LP decoding is a relaxed version of the Maximum-Likelihood (ML) decoding. The *codeword polytope* [3] of ML is replaced by a relaxed polytope, called the *fundamental polytope* [3]. The fundamental polytope arisen from a given parity check matrix. Its vertices are every codeword, but it also has some non-codeword. The vertices of the codeword polytope are the all codewords, and the vertices of the fundamental polytope are called *pseudocodewords* (PCWs) [3]. The additional non-codewords make the decision region [1] of the LP decoder even more complex than that of the ML. Therefore, a derivation of analytical bounds has an important role in evaluating the performance of the LP decoder.

The *fundamental cone* [2] is the conic hull of the fundamental polytope. The LP error probability over the fundamental polytope is equal to that over the fundamental cone [4]. Moreover, it is sufficient to consider only the fundamental cone *generators* [4] for evaluating the performance of the LP decoder.

The well-known upper bound on the error probability of a digital communication system is the *Union Bound* (UB), which is a first-order *Bonferroni-type* inequality [5] in the probability theory. The UB of the LP decoder [1] [6] [7] for High-Density Parity-Check (HDPC) codes presets inaccurate results due to the high density of fundamental cone generators. In fact, the union bound sums all of the pairwise error events as if they were disjoint, but this scenario is far from being the case in LP decoding of HDPC codes.

Each pseudocodeword in the LP decoder can be located in the BPSK signal space [2]. What the LP decoder does, it chooses the nearest pseudocodeword to the received vector as the most likely transmitted pseudocodeword. The ML soft decision decoder has such property as well, but

unlike to the LP decoder, its signal space contains only the set of the all codewords. Thus many of ML upper bounds can be reused [8] [9] [10] [11] in the case of LP decoding.

For a given code, each of its parity-check matrix creates a fundamental cone with different pseudo-weight spectrum and geometrical structure, which influences differently on the error probability of the LP decoder. Therefore, the geometrical properties of the fundamental cone generators are essential to evaluate with a better accuracy the LP decoding error probability. Thus ML error probability bounds which use the weight spectrum of the code or those who sum the error contribution of each individual codeword become less attractive. In [11] a ML bound is presented which is based on the second-order upper bound on the probability of a finite union of events. And indeed, it uses the geometrical properties of the codewords and considers an intersection of pairwise error events, but involves relatively high computational complexity.

To explore the density of the fundamental cone generators, we have defined the *angle graph*: each generator is considered as a node of a complete undirected graph. The cost of an edge is the angle between the generators related to the adjacent nodes. The minimum spanning tree is found and its cost distribution is illustrated. Different patterns for various parity-check matrices were observed.

In this paper, we propose an upper bound based on the second-order of Bonferroni-type inequality. The bound needs the fundamental cone generators rather than their weight spectrum. We call it Improved Linear Programming Union Bound (ILP-UB). It consists of two parts: The first term is the LP union bound itself, and the second term is a second-order correction that can be optimized by a known minimum spanning tree algorithm. It requires relatively low computational complexity since it involves only the Q-function.

The proposed ILP-UB makes use of an upper bound of the triplet-wise error probability that has been introduced earlier in the paper. We derive analytical expression to evaluate the triple-wise error probability depending on the angle which they create. And for example, the triple-wise error probability for the minimal-weight generators of the BCH[63,57,3] code is calculated. It is compared to the triple-wise error upper bound and to the UB in different angles and Signal-to-Noise Ratios (SNRs).

The proposed ILP-UB was tested on three HDPC codes: Golay[24,12,8], BCH[31,26,3], BCH[63 ,57,3], and on the Low-Density Parity-Check (LDPC) Tanner code [155,64,20] [12]. An improvement of up to 0.37 dB has been demonstrated over the conventional Linear Programming Union Bound (LP-UB).

This paper is organized as follows. Sec. II provides some background on ML and LP decoding. The minimum spanning tree problem for undirected graph is also reviewed in Sec. II. In Sec. III we explore the density of the fundamental cone generators and we check the effect of that density on the union bound of the triplet-wise error probability. The problem of finding an LP dominant error events is discussed in Sec. IV. In Sec. V we propose an improved linear programming error union bound. Sec. VI provides numerical results and discusses some possible direction for further research on how to improve the proposed bound. Sec. VII concludes the paper.

II. PRELIMINARIES AND DEFINITIONS

A. ML and LP Decoding

In this section we briefly review ML and LP decoding [3]. Consider a binary linear code $\mathcal{C}$ of length n, dimension k and code rate $R \triangleq k/n$. Let $\mathbb{F}_2 \triangleq \{0,1\}$ denote the finite field with two elements. The code $\mathcal{C}$ is defined by some $m \times n$ parity-check matrix $H \in \mathbb{F}_2^{m \times n}$ with row vectors $\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_m$, i.e. $\mathcal{C} \triangleq \{\mathbf{x} \in \mathbb{F}_2^n \mid \mathbf{x}H^T = 0\}$. The code will be called an [n,k,d] code, in which d is its minimum *Hamming distance*. The code is used for data communication over a memoryless binary-input channel with channel law $P_{Y|X}(y|x)$. We denote the transmitted codeword by $\mathbf{x} \triangleq (x_1, ..., x_n)$, the transmitted signal by $\overline{\mathbf{x}} \triangleq (\overline{x}_1, ..., \overline{x}_n)$ and the received signal by $\mathbf{y} \triangleq (y_1, ..., y_n)$. We assume that every codeword $\mathbf{x} \in \mathcal{C}$ is transmitted with equal probability. Let $\boldsymbol{\lambda}$ denote the Log-Likelihood Ratio (LLR) vector with the LLR components $\lambda_i \triangleq P_{Y|X}(y_i|0)/P_{Y|X}(y_i|1)$ for $i = 1, ..., n$. The block-wise Maximum Likelihood Decoding (MLD) is

$$\hat{\mathbf{x}}_{MLD}(\mathbf{y}) \triangleq \arg \min_{\mathbf{x} \in \mathcal{C}} \langle \mathbf{x}, \boldsymbol{\lambda} \rangle. \tag{1}$$

Where $\langle \mathbf{x}, \boldsymbol{\lambda} \rangle \triangleq \sum_i x_i \lambda_i$ denote the standard *inner product* of two vectors of equal length. The ML decoder error probability is independent of the transmitted CW, therefore, we assume without loss of generality that the all-zeros codeword $\mathbf{x}_0$ is transmitted. Then [13]

$$P_r^{MLD}(error \mid \mathbf{x}_0) = P_r\left(\hat{\mathbf{x}}_{MLD}(\mathbf{y}) \neq \mathbf{x}_0 \mid \mathbf{x}_0\right) \tag{2}$$

$$= P_r\left\{\bigcup_{x \in \mathcal{C} \backslash \mathbf{x}_0} ||\overline{\mathbf{x}} - \mathbf{y}||_2 \leq |||\overline{\mathbf{x}}_0 - \mathbf{y}||_2 \mid \mathbf{x}_0\right\} \tag{3}$$

$$\leq \sum_{x \in \mathcal{C} \backslash \mathbf{x}_0} P_r\left\{||\overline{\mathbf{x}} - \mathbf{y}||_2 \leq |||\overline{\mathbf{x}}_0 - \mathbf{y}||_2 \mid \mathbf{x}_0\right\} \tag{4}$$

$$= \sum_{x \in \mathcal{C} \backslash \mathbf{x}_0} Q\left(\frac{d_{\mathbf{x}}}{2\sigma}\right). \tag{5}$$

Where the Q-function is defined to be $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$ and $||\mathbf{x}||_2 \triangleq \sqrt{\sum_i x_i^2}$ denote the $\mathcal{L}_2$-norm of a vector $\mathbf{x}$. Eq. (3) also allows to make a simulation of the error probability contributed by a subgroup of codewords. Eq. (5) is the ML union bound, where $d_{\mathbf{x}} \triangleq ||\overline{\mathbf{x}} - \overline{\mathbf{x}}_0||_2 = 2\sqrt{RE_b w_H(\mathbf{x})}$ is the Euclidean distance from $\overline{\mathbf{x}}$ to the transmitted signal $\overline{\mathbf{x}}_0$.

The MLD (1) can be formulated [3] as the following equivalent optimization problem:

$$\hat{\mathbf{x}}_{MLD}(\mathbf{y}) \triangleq \arg\min_{\mathbf{x} \in conv(\mathcal{C})} \langle \mathbf{x}, \boldsymbol{\lambda} \rangle. \tag{6}$$

$conv(\mathcal{C})$ is called the *codeword polytope* [3], which is the convex hull of all possible codewords. The vertices of the codeword polytope are the all codewords. The number of inequalities needed to describe it grows exponentially in the code length. Therefore, solving this linear programming problem is not practical for codes with reasonable block length. To make this problem more feasible it was suggested [3] to replace $conv(\mathcal{C})$ by a relaxed polytope $\mathcal{P} \triangleq \mathcal{P}(H)$, called the *fundamental polytope.*

$$\mathcal{P} \triangleq \bigcap_{j=1}^{m} conv(\mathcal{C}_j) \quad \text{with} \quad \mathcal{C}_j \triangleq \left\{\mathbf{x} \in \mathbb{F}_2^n \mid \mathbf{x}\mathbf{h}_j^T = 0\right\}.$$

Where $conv(\mathcal{C}) \subseteq conv(\mathcal{C}_j)$ for $j = 1, ..., m$ and hence $conv(\mathcal{C}) \subseteq \mathcal{P}(H) \subset [0,1]^n$. The number of inequalities that describe $\mathcal{P}(H)$ is typically much smaller than those of $conv(\mathcal{C})$. The Linear Programming Decoding (LPD) is then

$$\hat{\boldsymbol{\omega}}_{LPD}(\mathbf{y}) \triangleq \arg\min_{\boldsymbol{\omega} \in \mathcal{P}} \langle \boldsymbol{\omega}, \boldsymbol{\lambda} \rangle. \tag{7}$$

In the case of $conv(\mathcal{C}) = \mathcal{P}(H)$ the relaxed LP solution equals to that of ML. In the case of

$conv(\mathcal{C}) \subset \mathcal{P}(H)$ the relaxed LP problem represents a suboptimal decoder which has vertices in $\mathcal{P}(H)$ which are not in $conv(\mathcal{C})$. The vertices of $\mathcal{P}(H)$, denoted by $\mathcal{V}(\mathcal{P}(H))$, are called LP pseudocodewords.

The *fundamental cone* [2] $\mathcal{K}(H) \triangleq \mathcal{K}$ is defined to be the conic hull of the fundamental polytope i.e. the set that consists of all possible conic combinations of all the points in $\mathcal{P}(H)$ and hence $\mathcal{P}(H) \subset \mathcal{K}(H)$. The LP decoding error probability over the fundamental polytope is equal to that over the fundamental cone [4]. We let $\mathbb{R}$ and $\mathbb{R}_+$ be the set of real numbers and the set of non-negative real numbers, respectively.

Definition 1. ([4], [14]) A set $\mathcal{G}(\mathcal{K}) \triangleq \{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_M \mid \mathbf{g}_i \in \mathbb{R}_+^n, \ i = 1, ..., M\}$ of M linearly independent vectors where $\mathcal{K} = \left\{ \sum_{i=1}^{M} \alpha_i \mathbf{g}_i \mid \alpha_i \in \mathbb{R} \right\}$ are called the *generators* of the cone $\mathcal{K}$. $\square$

It follows from Def. 1 that a vector $\mathbf{x}$ is in $\mathcal{K}$ if and only if $\mathbf{x}$ can be written as a nonnegative linear combination of the generators, i.e. $\mathbf{x} = \sum_{i=1}^{M} \alpha_i \mathbf{g}_i$ where $\alpha_i \in \mathbb{R}$. Note that a set of generators is not unique, and that the all-zeros codeword $\mathbf{x}_0 \notin \mathcal{G}(\mathcal{K})$.

We assume an AWGN channel, where each i-th transmitted bit perturbed by a white Gaussian noise z_i with a zero mean and noise power $\sigma^2 \triangleq N_0/2$. The received signal is $\mathbf{y} = \overline{\mathbf{x}} + \mathbf{z}$, where $\mathbf{z}$ designates an n-dimensional Gaussian noise vector with independent components $z_1, z_2, ..., z_n$.

We consider a BPSK modulation: the transmitted signal is $\overline{\mathbf{x}} = \gamma(1 - 2\mathbf{x})$, where $\gamma \triangleq \sqrt{RE_b}$ in which E_b is the information bit energy. The signal-to-noise ratio is defined to be $\mathrm{SNR} \triangleq E_b/N_0$. Following from the above, the LLR vector is $\boldsymbol{\lambda} = 4\frac{\sqrt{RE_b}}{N_0}\mathbf{y}$ [2], and therefore, the LPD will be considered henceforth

$$\hat{\boldsymbol{\omega}}_{LPD} = \arg\min_{\boldsymbol{\omega} \in \mathcal{P}} \langle \boldsymbol{\omega}, \mathbf{y} \rangle. \tag{8}$$

Definition 2. ([2], [15], [16]) Let $\boldsymbol{\omega} \in \mathbb{R}_+^n$. The AWGN channel pseudo-weight $w_p^{AWGNC}(\boldsymbol{\omega})$ of $\boldsymbol{\omega}$ is given by

$$w_p^{AWGNC}(\boldsymbol{\omega}) \triangleq \frac{||\boldsymbol{\omega}||_1^2}{||\boldsymbol{\omega}||_2^2}, \tag{9}$$

where $||\mathbf{x}||_1 \triangleq \sum_i |x_i|$ denote the $\mathcal{L}_1$-norm of a vector x. If $\boldsymbol{\omega} = 0$ we define $w_p^{AWGNC}(\boldsymbol{\omega}) \triangleq 0$, and in the case of $\boldsymbol{\omega} \in \{0, 1\}^n$ we have $w_p^{AWGNC}(\boldsymbol{\omega}) = w_H(\boldsymbol{\omega})$. $\square$

For an easier notation, as we discuss in this paper only AWGN channel, we will use the shorter notation $w_p(\boldsymbol{\omega})$ instead of $w_p^{AWGNC}(\boldsymbol{\omega})$.

Due to the symmetry property of the fundamental polytope the probability that the LP decoder fails is independent of the codeword that was transmitted [3]. Therefore, we henceforth assume without loss of generality when analyzing LPD error probability, that the all-zeros codeword $\mathbf{x}_0$ is transmitted.

The set of optimal solutions of a closed convex LP problem always includes at least one vertex of the polytope. Therefore, the LPD error probability is

$$P_r^{LPD}(error \mid \mathbf{x}_0) = P_r \left\{ \bigcup_{\boldsymbol{\omega} \in \mathcal{V}(\mathcal{P}(H)) \backslash \mathbf{x}_0} \langle \boldsymbol{\omega}, \mathbf{y} \rangle \leq 0 \mid \mathbf{x}_0 \right\}. \tag{10}$$

A pseudocodeword $\mathbf{p} \in \mathcal{V}(\mathcal{P})$ also belongs to the fundamental cone. Thus it can be written as a non-negative linear combination of the generators, i.e. $\mathbf{p} = \sum\limits_{i=1}^{M} \alpha_i \mathbf{g}_i$ with $\alpha_i \geq 0$. Therefore, if there is $\mathbf{p} \in \mathcal{V}(\mathcal{P})$ such that $\langle \mathbf{p}, \mathbf{y} \rangle = \sum\limits_{i=1}^{M} \alpha_i \langle \mathbf{g}_i, \mathbf{y} \rangle < 0$, then there must be at least one generator $\mathbf{g}_i \in \mathcal{G}(\mathcal{K})$ such that $\langle \mathbf{g}_i, \mathbf{y} \rangle < 0$. Therefore, the union of the pseudocodewords' error events in (10) can be replaced by the union of the generators' error events.

A vector $\boldsymbol{\omega} \in \mathbb{R}_+^n$ which is not codeword can be located into the signal space in the same way as a codeword, i.e $\overline{\boldsymbol{\omega}} = \gamma(1 - 2\boldsymbol{\omega})$. The vector $\boldsymbol{\omega}_{virt} \triangleq \frac{||\boldsymbol{\omega}||_1}{||\boldsymbol{\omega}||_2^2} \boldsymbol{\omega}$ was introduced by Vontobel and Koetter [2]. They showed that the decision hyperplane of $\boldsymbol{\omega}$ in the signal space, is at the same Euclidean distance from $\overline{\mathbf{x}}_0$ and from $\overline{\boldsymbol{\omega}}_{virt}$. Note that if $\boldsymbol{\omega} \in \mathcal{C} \subseteq \{0, 1\}^n$, then $\boldsymbol{\omega}_{virt} = \boldsymbol{\omega}$. From the above, the LP error probability is then expressed in the signal space as follows.

$$P_r^{LPD}(error \mid \mathbf{x}_0) = P_r \left\{ \bigcup_{\boldsymbol{\omega} \in \mathcal{G}(\mathcal{K}(H))} ||\overline{\boldsymbol{\omega}}_{virt} - \mathbf{y}||_2 \leq ||\overline{\mathbf{x}}_0 - \mathbf{y}||_2 \mid \mathbf{x}_0 \right\}. \tag{11}$$

Evaluating the LP error probability by simulating Eq. (11) is not practical, since it involves enormous number of generators. However, it allows to make a simulation of the error probability contributed by a subgroup of generators.

Let $E_{\mathbf{x}_0 \to \boldsymbol{\omega}} = \left\{ ||\overline{\boldsymbol{\omega}}_{virt} - \mathbf{y}||_2 \leq ||\overline{\mathbf{x}}_0 - \mathbf{y}||_2 \mid \mathbf{x}_0 \right\}$ denote the LP pairwise error event where the received vector $\mathbf{y}$ is closer to $\overline{\boldsymbol{\omega}}_{virt}$ than to the transmitted signal $\overline{\mathbf{x}}_0$. Thus the LP error

probability (11) can be written:

$$P_r^{LPD}(error \mid \mathbf{x}_0) = P_r \left\{ \bigcup_{\boldsymbol{\omega} \in \mathcal{G}(\mathcal{K}(H))} E_{\mathbf{x}_0 \to \boldsymbol{\omega}} \right\}, \tag{12}$$

and the LP union bound is

$$P_r^{LPD}(error \mid \mathbf{x}_0) \leq \sum_{\boldsymbol{\omega} \in \mathcal{G}(\mathcal{K}(H))} P_r\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}}\}. \tag{13}$$

Let $r_{\boldsymbol{\omega}} \triangleq \frac{\|\overline{\boldsymbol{\omega}}_{virt} - \overline{\mathbf{x}}_0\|_2}{2} = \gamma \sqrt{w_p(\boldsymbol{\omega})}$ denote the Euclidean distance from $\overline{\mathbf{x}}_0$ or from $\overline{\boldsymbol{\omega}}_{virt}$ to the decision boundary line. Thus the LP pairwise error probability [2]

$$P_r(E_{\mathbf{x}_0 \to \boldsymbol{\omega}}) = Q\left(\frac{r_{\boldsymbol{\omega}}}{\sigma}\right), \tag{14}$$

and the LP-UB in Eq. (13) can be written as follows [1] [7].

$$P_r^{LPD}(error \mid \mathbf{x}_0) \leq \sum_{\boldsymbol{\omega} \in \mathcal{G}(\mathcal{K}(H))} Q\left(\frac{r_{\boldsymbol{\omega}}}{\sigma}\right). \tag{15}$$

B. Undirected Graphs

In this section, we give a brief overview of some terms from graph theory. By a graph we will always mean an undirected graph without loops and multiple edges. We let $|V|$ denote the size of a set V.

Definition 3. ([17]) An *undirected graph* $G(V, \mathcal{E})$ consists of a set of nodes V and a set of edges $\mathcal{E}$. An edge is an unordered pair of nodes (v_i, v_j). Associated with each edge $(v_i, v_j) \in \mathcal{E}$ is a cost $c(v_i, v_j)$. □

Definition 4. ([17]) A *spanning tree* of an undirected graph $G(V, \mathcal{E})$, is a subgraph $T(V, \mathcal{E}')$ that is a tree and connects all the nodes in V. It has $|V|$ nodes and $|\mathcal{E}'| = |V| - 1$ edges, in which $\mathcal{E}'$ is a subset of $\mathcal{E}$. The cost of a spanning tree T, denoted by $cost(T)$, is the sum of the costs of all the edges in the tree. i.e. $cost(T) = \sum_{(v_i, v_j) \in T} c(v_i, v_j)$. □

Definition 5. ([17]) A spanning tree of a graph $G(V, \mathcal{E})$ is called a *Minimum Spanning Tree* (MST), if its cost is less than or equal to the cost of every other spanning tree $T(V, \mathcal{E}')$ of $G(V, \mathcal{E})$. □

Two popular algorithms for finding an MST in undirected graph are Prim's [18] and Kruskal's [19]. A simple implementation of Prim's algorithm can shows $O(|V|^2)$ running time, and both can be implemented with complexity of $O(|\mathcal{E}| \, log|V|)$.

III. GENERATOR DENSITY CHARACTERIZATION

In this section, we explore the density of the fundamental cone generators and we compare it to that of ML codewords. As a result, we will later examine how the union bound is affected by that density. Let $0 \leq \theta_{ij} \leq \pi$ denote the positive angle formed by the vectors $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$, which is equal to the angle formed by the vectors $\overrightarrow{\overline{\mathbf{x}}_0 \overline{\boldsymbol{\omega}}_{i,virt}}$ and $\overrightarrow{\overline{\mathbf{x}}_0 \overline{\boldsymbol{\omega}}_{j,virt}}$ in a BPSK signal space.

Definition 6. Let $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, ..., \boldsymbol{\omega}_M \in \mathbb{R}_+^n$ be a set of vectors. Consider each vector as a node of an undirected graph $G(V, \mathcal{E})$, with an undirected edge joining each pair of nodes $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$, denoted by $(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)$. An edge $(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) \in \mathcal{E}$ has a cost that equal to the angle between the vectors related to the adjacent nodes, i.e, $c(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) = \theta_{ij}$. The graph $G(V, \mathcal{E})$ will be called the *angle graph*. Note that the angle graph is a *complete* graph; it has $|V|$ nodes and $|V|(|V| - 1)/2$ edges. □

Definition 7. Let $T(V, \mathcal{E}')$ be an MST of the angle graph $G(V, \mathcal{E})$ in Def. 6. The *MST angle distribution* is defined to be the cost distribution of the all edges $(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)$ in the graph $T(V, \mathcal{E}')$. For easier notation, we will use the shorter term *angle distribution* instead. □

Example 8. Let $H_{G'}$ [1] and $H_{G''}$ (16) be parity-check matrices for the extended Golay[24,12,8] code. The former matrix was introduced by Halford and Chugg [20], the latter is a systematic parity-check matrix. Fig. 1 presents the angle distributions of the first 759 minimal-weight generators of $H_{G'}$ and $H_{G''}$ (generators with equal pseudo-weight were ordered randomly.). For a comparison, the angle distribution of the 759 minimal-weigh ML codewords is presented as well. The average angle of $H_{G'}$, $H_{G''}$ generators and of ML codewords are : $1.43°$, $10.69°$ and $60°$, respectively; and their Standard Deviations (STDs) are: $3.38°$, $8.72°$ and $0°$, respectively. Note that $H_{G'}$ and $H_{G''}$ have two different *generator matrices*, however, both have the same angle distribution for their 759 minimal-weight CWs. It is clear from Fig. 1, that $H_{G'}$ generators are much crowded than those of $H_{G''}$, and between these three distributions the ML codewords are spread most widely and evenly in the Euclidean space.

$$
H_{G''} = \begin{pmatrix}
0\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
1\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\
1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0 \\
1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0 \\
1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0 \\
1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0 \\
1\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0 \\
1\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0 \\
1\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1
\end{pmatrix}
\tag{16}
$$

Fig. 1. Angle distributions for the extended Golay[24, 12, 8] code of the first 759 minimal-weight generators of the parity-check matrices $H_{G'}$ and $H_{G''}$, compared to the angle distribution of the 759 minimal-weight ML codewords.

□

Example 9. The error probability contributed by two vectors depends on the angle between them. Let $\boldsymbol{\omega}_i, \boldsymbol{\omega}_j \in \mathbb{R}_+^n$ be vectors with an equal pseudo-weight, and let ξ_1 and ξ_2 be the two independent Gaussian random variables obtained by projecting the noise vector $\mathbf{z}$ onto the plan determined by the vectors $\overrightarrow{\overline{\mathbf{x}_0}\,\overline{\boldsymbol{\omega}}_{i,virt}}$ and $\overrightarrow{\overline{\mathbf{x}_0}\,\overline{\boldsymbol{\omega}}_{j,virt}}$. We refer to the probability $P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcup E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\}$ as the triplet-wise error probability, that is $\boldsymbol{\omega}_i$ or $\boldsymbol{\omega}_j$ was decoded when the all-zeros codeword was transmitted. The triplet-wise error probability depends on the angle θ_{ij} and it can be obtained by integrating a two dimensional Gaussian distribution over the darkened regions $\mathcal{R}_1$ and $\mathcal{R}_2$ in Fig. 2 [21]. Without loss of generality, we assume that $\boldsymbol{\omega}_j$ is placed on ξ_1 axis. $r_{\boldsymbol{\omega}_i}$ and $r_{\boldsymbol{\omega}_j}$ denote the Euclidean distances from the decision boundaries lines of $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$, respectively, to

the all-zeros codeword. In the case of vectors of equal pseudo-weight, $r_{\boldsymbol{\omega}_i} = r_{\boldsymbol{\omega}_j}$. The decision region boundary lines of $\boldsymbol{\omega}_i$ and $\boldsymbol{\omega}_j$ are $\xi_2 = -a\xi_1 + b$ and $\xi_1 = r_{\boldsymbol{\omega}_j}$, respectively. The $\boldsymbol{\omega}_i$ boundary line crosses ξ_2 axis at point $b = r_{\boldsymbol{\omega}_i}/sin\theta_{ij}$ and its slope is $a = \tan(90 - \theta_{ij})$. The intersection between the two boundary lines occurs at point $(\xi_1', \xi_2') = (r_{\boldsymbol{\omega}_j}, -ar_{\boldsymbol{\omega}_j} + b)$.

There are various numerical integration ways [22] to evaluate the triplet-wise error probability. Another possibility, is to approximate it by sum of Q-functions as follows.

$$P_r\left\{E_{\mathbf{x}_0 \to \omega_i} \bigcup E_{\mathbf{x}_0 \to \omega_j}\right\} = P_r\{\mathcal{R}_1\} + P_r\{\mathcal{R}_2\} \approx Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) +$$

$$\sum_{k=0}^{\left\lfloor \frac{\xi_{1,max}}{\triangle \xi_1} \right\rfloor} \left[1 - Q\left(\frac{-a(\xi_1' + k \triangle \xi_1) + b}{\sigma}\right)\right] \left[Q\left(\frac{\xi_1' + k \triangle \xi_1}{\sigma}\right) - Q\left(\frac{\xi_1' + (k+1) \triangle \xi_1}{\sigma}\right)\right]. \tag{17}$$

$P_r\{\mathcal{R}_1\}$ is equal to an LP pairwise error probability (14). $P_r\{\mathcal{R}_2\}$ is calculated as follows. The region $\mathcal{R}_2$ is divided into rectangles of a width $\triangle\xi_1$ which are parallel to the ξ_2 axis, as shown in Fig. 2. Each rectangle starts from a point on the decision boundary line of $\boldsymbol{\omega}_i$ and goes to infinity in the opposite direction of ξ_2 axis. The multiplication inside the sum of Eq. (17) is the probability that the noise components ξ_1 and ξ_2 are within the k-th rectangle. Since a two dimensional Gaussian distribution converges to zero as ξ_1 goes to infinity, it will be sufficient to sum from $k = 0$ to a large k such as $\left\lfloor \frac{\xi_{1,max}}{\triangle\xi_1} \right\rfloor$, where all the rectangles are located on the left side of the line $\xi_2 = \xi_{1,max}$.
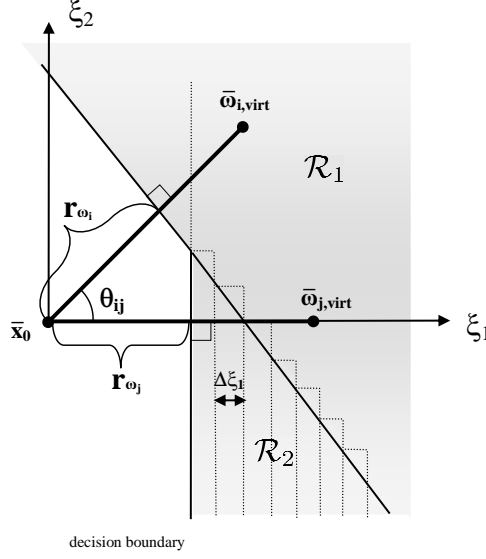
Fig. 2.   The LP triplet-wise error region in the signal space.

Example 10. Consider the BCH[63,57,3] code. The fundamental cone of the systematic parity-check matrix created by the *generator polynomial* x^6+x+1 has $11,551$ minimal-weight generators of pseudo-weight three. The angles between them varied from $5.85°$ to $90°$. The triplet-wise error probability of its two minimal-weight generators depends on θ_{ij} is presented in Fig. 3. It was calculated by Eq. (17) for 0 and 8 dB SNR in different angles. The triplet-wise union bound which is $2Q\left(\frac{r_\omega}{\sigma}\right)$ is presented as well. $\xi_{1,max}$ and $\triangle\xi_1$ was chosen to be 2000 and 1/2000, respectively. From Fig. 3 one can observe that the lower the SNR and the smaller the angle are, the worse is the UB. The figure also presents a triplet-wise error probability upper bound which is tighter than the UB and it will be introduced in Sec. V

(a) SNR = 0 dB

(b) SNR = 8 dB
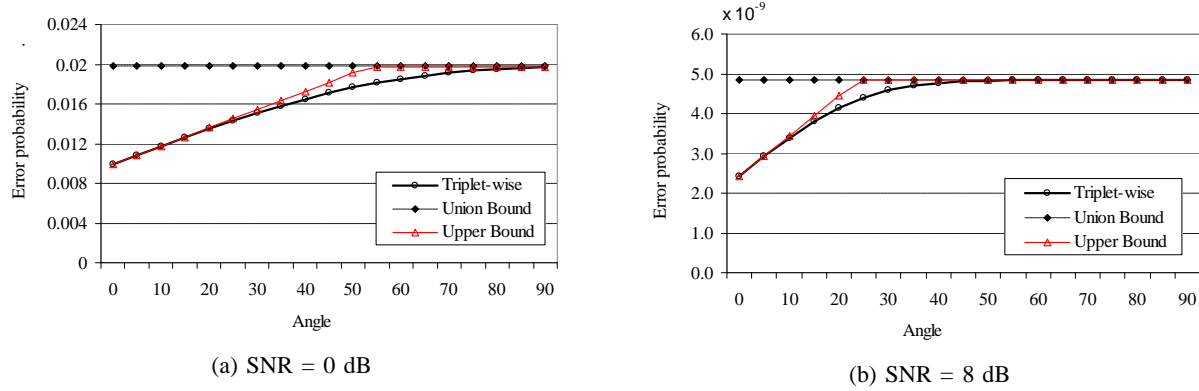
Fig. 3. Comparison between the triplet-wise error probability, its union bound and the upper bound in different angles of two minimal-weight generators of BCH[63,57,3], when the all-zeros word was transmitted.

□

IV. THE PROBLEM OF LOCATING DOMINANT ERROR EVENTS OF LPD

Consider a ML decoding of a binary-linear code BPSK-modulated over an AWGN channel. The decoder performance can be evaluated by considering the contributions of the most dominant error events to the probability of error. That dominant error events, especially in the higher SNR region, are the minimal weight codewords.

In this section, we will examine whether the minimal-weight generators of LP decoding have such a property as well. We let $w_H(\mathbf{x})$ denote the *Hamming weight* of $\mathbf{x}$, which is the number of non-zero positions of $\mathbf{x}$. Let $w_H^{min}(\mathcal{C})$ denote the minimum Hamming weight of a linear code $\mathcal{C}$, and let $w_p^{min}(H)$ denote the minimum AWGN channel pseudo-weight of a linear code defined by the parity-check matrix H. We will use the shorter notations w_H^{min} and w_p^{min} in case where the discussed code and matrix are mentioned explicitly. We let $\mathcal{K}_{\text{sub}} \subset \mathcal{K}$ denote a sub-cone of the fundamental cone which created by a chosen subgroup of generators. The LPD($\mathcal{K}_{\text{sub}}$) Frame Error Rate (FER) can be obtained by simulating Eq. (11). In the next example, we will study the error probability contributed by a subgroup of codewords and generators for the extended Golay[24,12,8] code.

Example 11. The extended Golay[24,12,8] code has a total 4,096 codewords of which 759 have minimal Hamming weight of $w_H^{min} = 8$. The fundamental cone of the parity check-matrix $H_{G'}$

has a total of $231,146,333$ generators of which two have minimal-weight of $w_p^{min} = 3.6$ [1]. Simulating the error probability by Eq. (3) shows that the minimal-weight CWs describe well the MLD performance at the whole range of SNR, which is not the case for the first 759 minimal-weight generators for LPD. For instance, consider the error rate of 10^{-2}, it was found that the difference between LPD($\mathcal{K}_{sub}$) and LPD($\mathcal{K}$) is about 2.5 dB. The angle distributions which were presented in Fig. 1 support this result: the average angle of that group of generators is as small as $1.43°$, and the average angle of the ML minimal-weight CWs is $60°$. $\qquad\square$

There are number of reasons why the minimal-weight generators are often not a dominant subgroup of LPD: (a) There is no guarantee for significant number of generators with minimal pseudo-weight. The fundamental cone of $H_{G'}$ for example, has only two. (b) A subgroup of generators can be very crowded, which significantly reduces their contribution to the error probability. (c) Unlike MLD which has distinct subgroup of minimal-weight codewords, LPD often has a continuous-like weight distribution. For example, the BCH[31,21,5] code of parity-check matrix $H_{BCH_{[31,21]}}$ (18) has 627,052,479 generators. The pseudo-weight distribution of these generators is presented in Fig. 4. Its smooth distribution makes it difficult to locate a minimal-weight dominant subgroup.

In LPD, a potential subgroup to be a dominant is taking all generators of weight $w_p \leq w_H^{min}$. This group is not empty since $w_p^{min} \leq w_H^{min}$ [23], however, it may contains enormous number of generators. For example, Golay[24,12,8] has only 759 minimal-weight CWs of $w_H^{min} = 8$, but the fundamental cone of parity-check matrix $H_{G''}$ has $143,757,418$ generators of weight $w_p \leq w_H^{min} = 8$.

$$H_{BCH_{[31,21]}} = \begin{pmatrix} 1000000000011010101111001001010\\ 0100000000011010101111001001010\\ 0010000000011010101111001001010\\ 0001000000011001111010110000110\\ 0000100000011001111010110000011\\ 0000010000111001100011001010101\\ 0000001000101001101111110111110\\ 0000000100010100110111111011111\\ 0000000010101111100010110111011\\ 0000000001101010111100100101001 \end{pmatrix} \qquad (18)$$
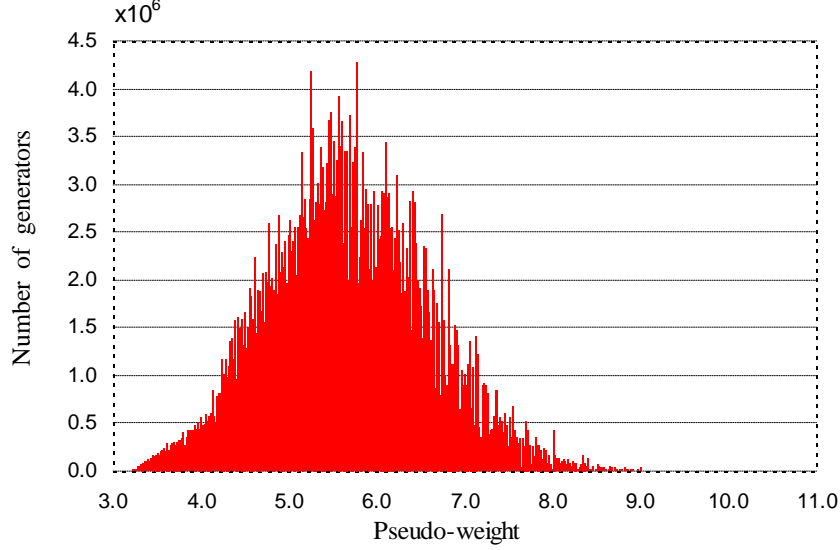
Fig. 4. A complete generators' pseudo-weight distribution for the BCH[31,21,5] code of $H_{BCH_{[31,21]}}$ with 627,052,479 generators.

V. IMPROVED LP UNION BOUND

In this section, we propose an improved union bound for LP decoding of a binary linear code transmitted over a binary-input AWGN channel. This bound is based on the second-order of Bonferroni-type inequality in probability theory [5], also referred to as Hunter bound [24]. For any set of events $E_1, E_2, ..., E_M$ and their *complementary* events, denoted by $E_1^c, E_2^c, ..., E_M^c$,

$$P_r\left(\bigcup_{i=1}^{M} E_i\right) = \sum_{i=1}^{M} P_r\left(E_i \bigcap \left[\bigcup_{j=1}^{i-1} E_j^c\right]\right). \tag{19}$$

Let denote the $M!$ possible permutations of the indices of the error events $E_1, E_2, ..., E_M$ by $\Pi(1,2,...,M) = \{\pi_1, \pi_2,...,\pi_M\}$. For a given Π, let $\Lambda = \{\hat{\pi}_2, \hat{\pi}_3, ..., \hat{\pi}_M\}$ denote the $(M^2 - M)/2$ possible sets of indices in which $\hat{\pi}_i \in \{\pi_1, \pi_2, ..., \pi_{i-1}\}$ for $i = 2, 3, ..., M$. Hunter [24] presented the second-order bound of Eq. (19) as follows.

$$P_r\left(\bigcup_{i=1}^{M} E_i\right) \leq \sum_{i=1}^{M} P_r(E_{\pi_i}) - \sum_{i=2}^{M} P_r(E_{\pi_i} \cap E_{\hat{\pi}_i}). \tag{20}$$

Minimization of the Right-Hand Side (RHS) of Eq. (20) is required to achieve the tightest second-order bound. Using the sets of the indices Λ and Π, the minimization problem can be

written as follows [10] [24].

$$P_r\left(\bigcup_{i=1}^{M}E_i\right) \le \sum_{i=1}^{M}P_r(E_i) + \min_{\Pi,\Lambda}\left\{-\sum_{i=2}^{M}P_r(E_{\pi_i}\cap E_{\hat{\pi}_i})\right\}. \tag{21}$$

The first sum goes through over all the indices 1 to M of the error events, thus E_{π_i} could be changed to E_i.

Consider each of the random events E_i as a node of an undirected graph G and the intersection $(E_i \cap E_j)$ as an undirected edge joining the nodes E_i and E_j, denoted by (i,j), with a cost $c(i,j) = P_r(E_i \cap E_j)$. Hunter [24] showed that a set of $(M-1)$ intersections may be used in the second term of Eq. (21) if and only if it forms a spanning tree of the nodes $\{E_i\}_{i=1}^{M}$. Thus the minimization problem of Eq. (21) can be written equivalently [24], [10],

$$P_r\left(\bigcup_{i=1}^{M}E_i\right) \le \sum_{i=1}^{M}P_r(E_i) + \min_{\tau}\left\{-\sum_{(i,j)\in\tau}P_r(E_i\cap E_j)\right\}. \tag{22}$$

Where τ is a spanning tree of the graph G. The problem is to find a graph τ which minimizes Eq. (22) over all possible spanning trees. The solution for that is known as the solution of the minimum spanning tree problem and has been proposed by Prim [18] and Kruskal [19].

Consider the event E_i as the pairwise error event $E_{\mathbf{x}_0\to\boldsymbol{\omega}_i}$. In order to upper bound the LP decoding error probability in Eq. (12) by the second-order upper bound (22), the probability $P_r\left\{E_{\mathbf{x}_0\to\boldsymbol{\omega}_i}\bigcap E_{\mathbf{x}_0\to\boldsymbol{\omega}_j}\right\}$ is required, or instead, its lower bound. The probability of intersection of two events can be expressed using the *inclusion-exclusion* principle in probability theory,

$$P_r\left\{E_{\mathbf{x}_0\to\boldsymbol{\omega}_i}\bigcap E_{\mathbf{x}_0\to\boldsymbol{\omega}_j}\right\} = P_r\left\{E_{\mathbf{x}_0\to\boldsymbol{\omega}_i}\right\} + P_r\left\{E_{\mathbf{x}_0\to\boldsymbol{\omega}_j}\right\} - P_r\left\{E_{\mathbf{x}_0\to\boldsymbol{\omega}_i}\bigcup E_{\mathbf{x}_0\to\boldsymbol{\omega}_j}\right\}. \tag{23}$$

The first and the second terms in the RHS of Eq. (23) are the LP pairwise error probability (14), the third term can be upper bounded by the following theorem.

Theorem 12. *Let $\boldsymbol{\omega}_i, \boldsymbol{\omega}_j \in \mathbb{R}_+^n$ be vectors of a pseudo-weight $w_p(\boldsymbol{\omega}_i) \ne w_p(\boldsymbol{\omega}_j)$. The LP triplet-wise error probability*

$$P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcup E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\} \le min \left\{ \begin{array}{c} Q\left(\frac{min(r_{\boldsymbol{\omega}_i}, r_{\boldsymbol{\omega}_j})}{\sigma}\right) + \frac{\theta_{ij}}{2\pi}e^{-\frac{max(r_{\boldsymbol{\omega}_i}^2, r_{\boldsymbol{\omega}_j}^2)}{2\sigma^2}}, \\ Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) + Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) - Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right)Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) \end{array} \right\}. \tag{24}$$

Proof: Let $\tilde{\xi} \triangleq \xi_1^2 + \xi_2^2$ be a random variable with Chi-square distribution [25] with two degrees of freedom, i.e.

$$f(\tilde{\xi}) = \frac{1}{2\sigma^2}e^{-\frac{\tilde{\xi}}{2\sigma^2}}U(\tilde{\xi}), \tag{25}$$

in which $U(\cdot)$ is the unit step function. Without loss of generality we assume that $w_p(\boldsymbol{\omega}_i) < w_p(\boldsymbol{\omega}_j)$. With the help of Fig. 5 the triplet-wise error probability,

$$P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcup E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\} \le P_r\left(\bigcup_{i=1}^4 \mathcal{R}_i\right) \le \sum_{i=1}^4 P_r\{\mathcal{R}_i\} \tag{26}$$

$$= \underbrace{Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right)}_{P_r(\mathcal{R}_1)+P_r(\mathcal{R}_2)} + \underbrace{\frac{\theta_{ij}}{2\pi}P_r\left(\tilde{\xi} > r_{\boldsymbol{\omega}_j}^2\right)}_{P_r(\mathcal{R}_3)} + \underbrace{Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) - Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right)}_{P_r(\mathcal{R}_4)} \tag{27}$$

$$= Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) + \frac{\theta_{ij}}{2\pi}e^{-\frac{r_{\boldsymbol{\omega}_j}^2}{2\sigma^2}}. \tag{28}$$

From the noise symmetry, each of the probabilities $P_r(\mathcal{R}_1)$ or $P_r(\mathcal{R}_2)$ equal to $\frac{1}{2}Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right)$. $P_r(\mathcal{R}_3)$ is the probability that of $\xi_1^2 + \xi_2^2$ lies in the region outside a circle of a radios r_{ω_j} created by the central angle θ_{ij}. $P_r\left(\tilde{\xi} > r_{\boldsymbol{\omega}_j}^2\right)$ was calculated in Eq. (27) by integrating the Chi-square distribution (25) from $r_{\boldsymbol{\omega}_j}^2$ to ∞. Thus for two vectors of pseudo-weight $w_p(\boldsymbol{\omega}_i) \ne w_p(\boldsymbol{\omega}_j)$

$$P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcup E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\} \le Q\left(\frac{min(r_{\boldsymbol{\omega}_i}, r_{\boldsymbol{\omega}_j})}{\sigma}\right) + \frac{\theta_{ij}}{2\pi}e^{-\frac{max(r_{\boldsymbol{\omega}_i}^2, r_{\boldsymbol{\omega}_j}^2)}{2\sigma^2}}. \tag{29}$$

The triplet-wise error probability can also be bounded using the inclusion–exclusion principle as follows.

$$P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcup E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\} = P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i}\right\} + P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\} - P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcap E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\} \tag{30}$$

$$\le Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) + Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) - Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right)Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right). \tag{31}$$

The transition from Eq. (30) to Eq. (31) was done by lower bounding $P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcap E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\}$ at

its lowest value $Q\left(\frac{r_{\omega_i}}{\sigma}\right)Q\left(\frac{r_{\omega_j}}{\sigma}\right)$ accepted in $\theta_{ij} = 90^0$. Finally, selecting the minimum between Eq. (29) and Eq. (31) completes the proof. ∎
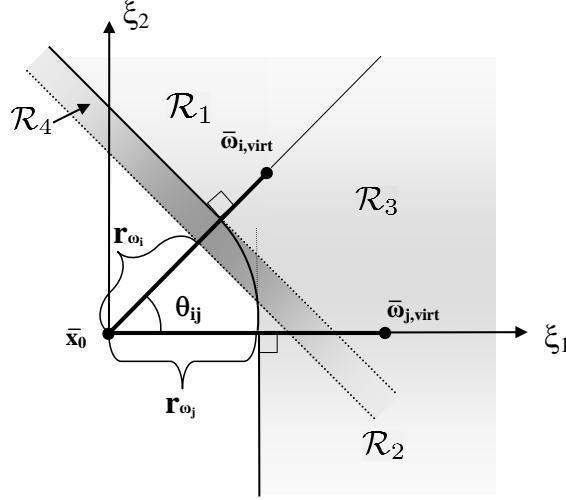

Fig. 5. The region in the signal space used to bound the LP triplet-wise error probability $(w_p(\boldsymbol{\omega}_i) \neq w_p(\boldsymbol{\omega}_j))$.

Example 13. We continue Ex. 10. The triplet-wise error probability upper bound of Theorem 12 was calculated for two minimal-weight generators of the BCH[63,57,3] code. It is presented in Fig. 3 together with the previous results of Ex. 10. We can see that the smaller the angle and lower the SNR, the more improvement the triplet-wise error upper bound has over the union bound. Note that because $\frac{r_{\boldsymbol{\omega}}}{\sigma} \propto \sqrt{\text{SNR} \cdot w_p(\boldsymbol{\omega})}$, changing the pseudo-weight of the generators will have the same effect as changing the SNR. Thus this bound is expected to have more improvement on low pseudo-weight generators. □

In the next theorem, we propose an improved UB for the LP decoding.

Theorem 14. *Let $\mathcal{G}(\mathcal{K}(H))$ be a set of cone generators of a parity-check matrix H. For each $\boldsymbol{\omega}_i \in \mathcal{G}$ the pairwise error event $E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i}$ is considered as a node of a complete graph $G(V, \mathcal{E})$. Let $(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j)$ denote an undirected edge joining the nodes related to the events $E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i}$ and $E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}$. $\tau(V, \mathcal{E}')$ is denoted for a spanning tree of $G(V, \mathcal{E})$. The LP decoding error probability can be upper-bounded by*

$$P_r^{LPD}(error \mid \mathbf{x}_0) \leq \sum_{\boldsymbol{\omega} \in \mathcal{G}(\mathcal{K}(H))} Q\left(\frac{r_{\boldsymbol{\omega}}}{\sigma}\right)$$

$$+ \min_{\tau} \left\{ \sum_{(\boldsymbol{\omega}_i, \boldsymbol{\omega}_j) \in \tau} \min \left\{ \begin{array}{c} -Q\left(\frac{max(r_{\boldsymbol{\omega}_i}, r_{\boldsymbol{\omega}_j})}{\sigma}\right) + \frac{\theta_{ij}}{2\pi} e^{-\frac{max(r_{\boldsymbol{\omega}_i}^2, r_{\boldsymbol{\omega}_j}^2)}{2\sigma^2}}, \\ -Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) \end{array} \right\} \right\} \quad (32)$$

We call this bound the *Improved LP Union Bound* (ILP-UB). The first term is the LP union bound itself (15), the second term is a second-order correction.

Proof: To prove this, we will apply Hunter bound for the LP error probability. First, we find a lower bound for $P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcap E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\}$: by substituting the upper bound of $P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcup E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\}$ (24) into the inclusion–exclusion principal (23), we will have

$$P_r\left\{E_{\mathbf{x}_0 \to \boldsymbol{\omega}_i} \bigcap E_{\mathbf{x}_0 \to \boldsymbol{\omega}_j}\right\} \geq$$

$$\geq Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) + Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) - \min \left\{ \begin{array}{c} Q\left(\frac{min(r_{\boldsymbol{\omega}_i}, r_{\boldsymbol{\omega}_j})}{\sigma}\right) + \frac{\theta_{ij}}{2\pi} e^{-\frac{max(r_{\boldsymbol{\omega}_i}^2, r_{\boldsymbol{\omega}_j}^2)}{2\sigma^2}}, \\ Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) + Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) - Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) \end{array} \right\} \quad (33)$$

$$= \max \left\{ \begin{array}{c} Q\left(\frac{max(r_{\boldsymbol{\omega}_i}, r_{\boldsymbol{\omega}_j})}{\sigma}\right) - \frac{\theta_{ij}}{2\pi} e^{-\frac{max(r_{\boldsymbol{\omega}_i}^2, r_{\boldsymbol{\omega}_j}^2)}{2\sigma^2}}, \\ Q\left(\frac{r_{\boldsymbol{\omega}_i}}{\sigma}\right) Q\left(\frac{r_{\boldsymbol{\omega}_j}}{\sigma}\right) \end{array} \right\}. \quad (34)$$

Applying Hunter bound (22) for LP decoding error probability (12) and substituting into it the expression in (34) together with the LP pairwise error probability (14), will give the desired result. ∎

Given a set of generators $\mathcal{G}$, the running time of ILP-UB is equal to that of finding an MST on a complete graph $G(V, \mathcal{E})$. It can be obtained by Prim's algorithm with a complexity of $O(|\mathcal{G}|^2)$. The LP-UB for a comparison, for a given set of generators has running time of $O(|\mathcal{G}|)$.

VI. RESULTS AND DISCUSSION

In this section, we provide results to show the improvement of ILP-UB over LP-UB. For this purpose, we examine four codes, three HDPC codes: extended Golay[24,12,8], BCH[31,26,3], BCH[63,57,3]; and one LDPC Tanner code [155,64,20] [12]. The parity-check matrices we use for Golay[24,12,8] and BCH[31,26,3] are H_G'' (16) and $H_{BCH_{[31,26]}}$ (35), respectively; and for the BCH[63,57,3] we use a systematic parity-check matrix created by the generator polynomial $x^6 + x + 1$. The minimal pseudo-weight of the extended Golay[24,12,8] is $w_p^{min} = 3.2$. BCH[31,26,3] and BCH[63,57,3] have the same minimal pseudo-weight: $w_p^{min} = 3$; and the Tanner code [155,64,20] has $w_p^{min} \approx 16.403$ [1].

Because of the enormous number of cone generators, we chose representative subgroups: for the BCH[31,26,3], BCH[63,57,3] and Tanner code [155,64,20] we chose all the minimal-weight generators that are 1,185 , 11,551 and 465 generators, respectively. Because the extended Golay[24,12,8] code has only 165 minimal-weight generators we chose for it the first 231 generators of a weight equal or less than $w_p = 3.25$.

$$H_{BCH_{[31,26]}} = \begin{pmatrix} 1\,0\,0\,0\,0\,1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,1\,1\,1\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,0\,1\,0 \\ 0\,1\,0\,0\,0\,0\,1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,1\,1\,1\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,0\,1 \\ 0\,0\,1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,1\,1\,1\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,0\,1\,0\,1\,0\,0 \\ 0\,0\,0\,1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,1\,1\,1\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,0\,1\,0\,1\,0 \\ 0\,0\,0\,0\,1\,0\,0\,1\,0\,1\,1\,0\,0\,1\,1\,1\,1\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,0\,1\,0\,1 \end{pmatrix} \tag{35}$$
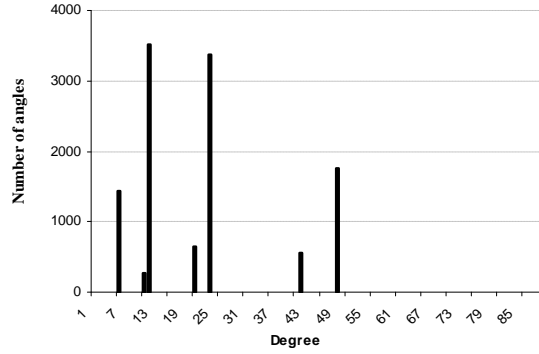
Fig. 6 presents the angle distributions according to Def. 7 for the aforementioned codes: extended Golay[24,12,8], BCH[31,26,3] and BCH[63,57,3]. Their average angles are $19.85°$, $29.58°$, $21.87°$, respectively; and their STDs are $13.44°$, $13.94°$, $13.84°$, respectively.
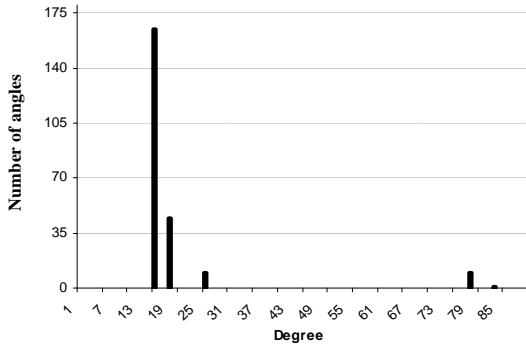
(a) Golay[24,12,8] code: angle distribution for all the 231 generators with $w_p \leq 3.25$.
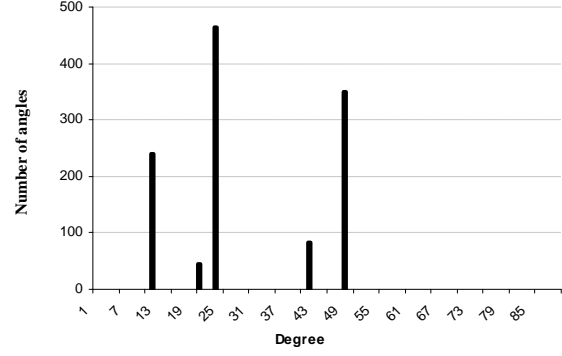
(b) BCH[31,26,3] code: angle distribution of all the 1,185 minimal-weight generators.


(c) BCH[63,57,3] code: angle distribution of all the 11,551 minimal-weight generators.

Fig. 6. Angle distributions.

Fig. 7 presents results of the: ILP-UB($\mathcal{K}_{\text{sub}}$), LP-UB($\mathcal{K}_{\text{sub}}$) and LPD($\mathcal{K}_{\text{sub}}$) for the chosen subgroups of generators. It presents the LPD FER as well. The ILP-UB optimized by Prim's algorithm. The ILP-UB presents an improvement over the LP-UB. For instance, we consider the error rate of 10^{-2}. For the extended Golay[24,12,8], the difference between LP-UB($\mathcal{K}_{\text{sub}}$) and LPD($\mathcal{K}_{\text{sub}}$) is about 0.9 dB while ILP-UB($\mathcal{K}_{\text{sub}}$) shows an improvement of 0.37 dB over LP-UB($\mathcal{K}_{\text{sub}}$). For BCH[31,26,3], the difference between LP-UB($\mathcal{K}_{\text{sub}}$) and LPD($\mathcal{K}_{\text{sub}}$) is about 0.47 dB while ILP-UB($\mathcal{K}_{\text{sub}}$) shows an improvement of 0.13 dB. And for BCH[63,57,3], the difference between LP-UB($\mathcal{K}_{\text{sub}}$) and LPD($\mathcal{K}_{\text{sub}}$) is about 0.62 dB while ILP-UB($\mathcal{K}_{\text{sub}}$) shows an improvement of 0.16 dB.

The results of the LDPC Tanner code were omitted, since the improvement of the ILP-UB($\mathcal{K}_{\text{sub}}$) over the LP-UB($\mathcal{K}_{\text{sub}}$) at error rate of 10^{-3} is dropped to about 0.05 dB. The reason for

that is twofold. First, the Tanner code has a large average angle: $35.16°$. Second, the generators have an high pseudo-weight: $w_p^{min} \approx 16.403$. These two values are high as compared to the other tested codes.
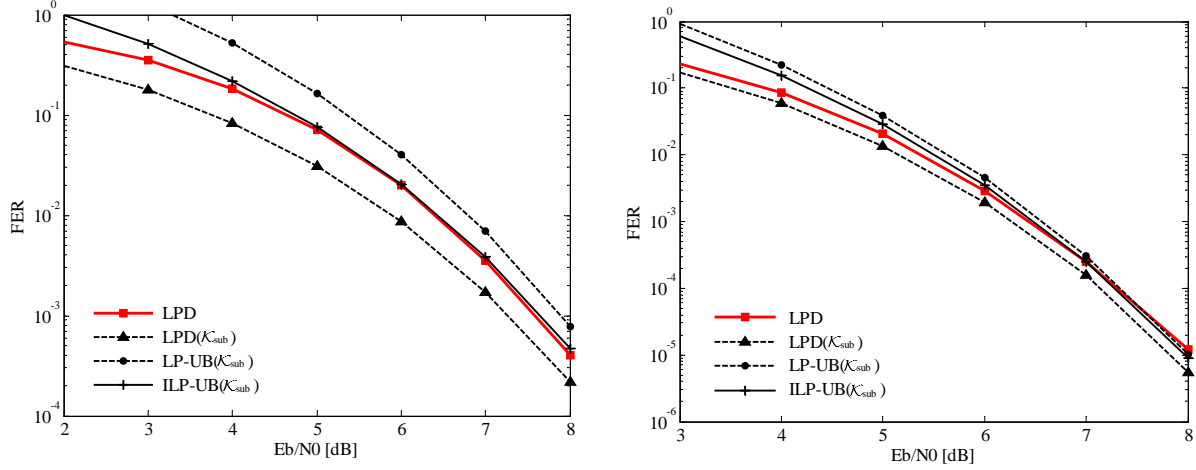
(a) Golay[24,12,8] code: results for 231 generators with $w_p \leq 3.25$ ($w_p^{min} = 3.2$).

(b) BCH[31,26,3] code: results of all the $1,185$ minimal-weight generators ($w_p^{min} = w_H^{min} = 3$).
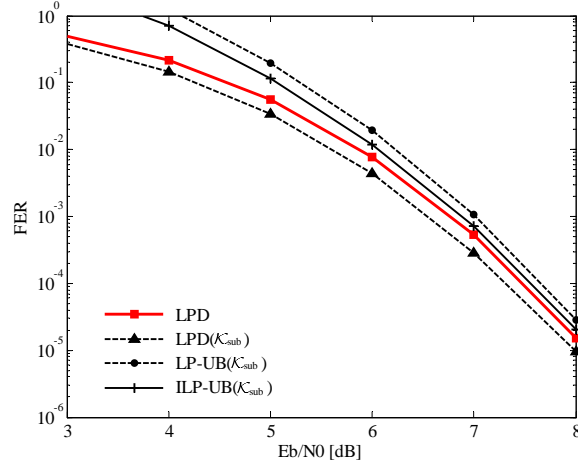
(c) BCH[63,57,3] code: results of all the $11,551$ minimal-weight generators ($w_p^{min} = w_H^{min} = 3$).

Fig. 7. A comparison between ILP-UB, LP-UB, LPD and LPD FER for HDPC codes.

Fig . 7 together with Fig. 6 show that the lower the average angle is, the more improvement the ILP-UB has. A small average angle is typical for HDPC codes, therefore, the advantage of ILP-UB over the LP-UB will be reflected better on such type of codes. But on the other hand, as the larger the average angle is, the better the LP-UB will be. Fig. 7a presents the highest improvement of the ILP-UB($\mathcal{K}_{sub}$) among the other codes. This result correlates to Golay's smallest average

angle: $19.85°$. However, it presents the largest gap to its $\text{LPD}(\mathcal{K}_{\text{sub}})$. This apparently happens because there are a significant probabilities of intersections between three error events or more.

Bukszár and Prékopa have suggested [26] a third order upper bound on the probability of a finite union of events. Their bound considers intersections of two and three events. They proved that this third order bound, which is obtained by the use of a type of graph called cherry tree, is at least as strong as the second-order bound. Therefore, implementing such a bound will improve (or at least will be equal to) the proposed ILP-UB.

VII. CONCLUSIONS

In this paper, we have presented an improved union bound on the error probability of LP decoding of binary linear HDPC codes transmitted over a binary-input AWGN channel. It is based on the second-order upper bound on the probability of a finite union of events. It has low computational complexity since it only involves the Q-function. It can be implemented with running time of $O(|\mathcal{G}|^2)$, where $\mathcal{G}$ is a set of generators of the fundamental cone arisen from a given parity check matrix. We examined the proposed bound for several HDPC codes: Golay[24,12,8], BCH[31,26,3], BCH[63,57,3], and for the LDPC Tanner [155,64,20] code. The improvement of the proposed bound over the union bound presents dependency on the pseudo-weight of the generators and their density. We studied and compared the generator density through the angle distribution of various codes and parity-check matrices. Finally, a third order upper bound was proposed, it is based on a type of graph called cherry tree, and is left open for further research.

REFERENCES

[1] A. Lifshitz and Y. Be'ery, "On pseudocodewords and decision regions of linear programming decoding of HDPC codes," *to appear in IEEE Trans. Commun.*, 2012. [Online]. Available: http://arxiv.org/abs/1105.2631

[2] P. O. Vontobel and R. Koetter, "Graph-cover decoding and finite-length analysis of message-passing iterative decoding of LDPC codes," 2005. [Online]. Available: http://www.arxiv.org/abs/cs.IT/0512078

[3] J. Feldman, M. Wainwright, and D. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 954–972, March 2005.

[4] P. Chaichanavong and P. Siegel, "Relaxation bounds on the minimum pseudo-weight of linear block codes," in *Proc. IEEE International Symposium on Inf. Theory*, September 2005, pp. 805–809. [Online]. Available: http://www.arxiv.org/abs/cs.IT/0508046

[5] J. Galambos and I. Simonelli, *Bonferroni-type Inequalities with Applications.* New York, NY: Springer, 1996.

[6] R. Smarandache and P. Vontobel, "Pseudo-codeword analysis of tanner graphs from projective and euclidean planes," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2376–2393, July 2007.

[7] V. Skachek and M. Flanagan, "Lower bounds on the minimum pseudodistance for linear codes with q-ary PSK modulation over AWGN," in *Proc. 5th International Symposium on Turbo Codes and Related Topics, Lausanne, Switzerland*, September 2008, pp. 426–431.

[8] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1284–1292, July 1994.

[9] H. Herzberg and G. Poltyrev, "Techniques of bounding the probability of decoding error for block coded modulation structures," *IEEE Trans. Inf. Theory*, vol. 40, no. 3, pp. 903–911, May 1994.

[10] S. Yousefi and A. Khandani, "A new upper bound on the ML decoding error probability of linear binary block codes in AWGN interference," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3026–3036, December 2004.

[11] A. Mehrabian and S. Yousefi, "Improved tangential sphere bound on the ML decoding error probability of linear binary block codes in AWGN and block fading channels," *IEE Proc. Commun.*, vol. 153, no. 6, pp. 885–893, December 2006.

[12] R. Tanner, D. Sridhara, A. Sridharan, T. Fuja, and J. Costello, D.J., "LDPC block and convolutional codes based on circulant matrices," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 2966–2984, December 2004.

[13] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication, 3rd ed.*, Springer, 2004.

[14] W. Kositwattanarerk and G. Matthews, "Lifting the fundamental cone and enumerating the pseudocodewords of a parity-check code," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 898–909, February 2011.

[15] N. Wiberg, "Codes and decoding on general graphs," Ph.D. dissertation, Linköping University, Linköping, Sweden, 1996.

[16] G. D. Forney, Jr., R. Koetter, F. R. Kschischang, and A. Reznik, "On the effective weights of pseudocodewords for codes defined on graphs with cycles," in *Codes, Systems, and Graphical Models, ser. IMA Vol. Math. Appl.* Springer, November 1998, pp. 101–112.

[17] A. V. Aho, J. E. Hopcroft, and J. Ullman, *Data Structures and Algorithms*. Boston, MA, USA: Addison-Wesley, 1983.

[18] R. C. Prim, "Shortest connection networks and some generalizations," *Bell Syst. Tech. J.*, vol. 36, pp. 1389–1401, November 1957.

[19] J. Kruskal, Joseph B., "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proc. Amer. Math. Soc*, vol. 7, no. 1, pp. 48–50, February 1956.

[20] T. Halford and K. Chugg, "Random redundant soft-in soft-out decoding of linear block codes," in *Proc. IEEE International Symposium on Inf. Theory, Seattle, Washington, USA*, July 2006, pp. 2230–2234.

[21] X. Ma, J. Liu, and B. Bai, "New techniques for upper-bounding the ML decoding performance of binary linear codes," 2011. [Online]. Available: http://arxiv.org/abs/1104.1471

[22] R. Burden and J. Faires, *Numerical Analysis*, 8th ed. Thomson Brooks/Cole, 2005.

[23] P. Vontobel and R. Koetter, "Lower bounds on the minimum pseudoweight of linear codes," in *Proc. IEEE International Symposium on Inf. Theory 2004*, June 2004, p. 70.

[24] D. Hunter, "An upper bound for the probability of a union," *J. Applied Probability*, vol. 13, no. 3, pp. 597–603, 1976.

[25] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*. New York, NY: John Wiley & Sons, 1994, vol. 1.

[26] J. Bukszár and A. Prékopa, "Probability bounds with cherry trees," *Math. Oper. Res.*, vol. 26, no. 1, pp. 174–192, 2001.